**Data Collection and Preprocessing Phase**

|  |  |
| --- | --- |
| Date | 15 June 2024 |
| Team ID | 739820 |
| Project Title | Predicting the Unpredictable: A Look into the World of powerlifting |
| Maximum Marks | 2 Marks |

**Data Quality Report:**

The Data Quality Report will summarize data quality issues from the selected source, including severity levels and resolution plans. It will aid in systematically identifying and rectifying data discrepancies.

**Data Quality Report:**

|  |  |  |  |
| --- | --- | --- | --- |
| **Data Source** | **Data Quality Issue** | **Severity** | **Resolution Plan** |
| Competition Results Database | Missing values in ‘Squat’,’Bench Press’,’Deadlift’,and ‘Total’  columns | High | Use mean/median imputation or expert estimation based on other available data |
| Athlete Training Logs | Incomplete training logs due to inconsistent data entry | High | Implement mandatory data entry protocols and use interpolation for missing data |